MIMEOGRAPH SERIES #1629

GOODNESS OF FIT TESTS FOR OPEN CAPTURE-RECAPTURE MODELS

Kenneth H. Pollock

Department of Statistics, North Carolina State University
Box 5457, Raleigh, North Carolina 27650, U.S.A.

and

James E. Hines

James D. Nichols

Patuxent Wildlife Research Center, U.S. Fish and Wildlife Service
Laurel, Maryland 20811, U.S.A.

SUMMARY

General goodness of fit tests for the Jolly-Seber model are proposed. These tests are based on conditional arguments using minimal sufficient statistics. The tests are shown to be of simple hypergeometric form so that a series of independent contingency table chi-square tests can be performed. The relationship of these tests to other proposed tests is discussed. This is followed by a simulation study of the power of the tests to detect departures from the assumptions of the Jolly-Seber model.

Key words: Capture-recapture; Jolly-Seber model; Goodness of fit tests;

1. Introduction

Jolly (1965) and Seber (1965) develop a stochastic k-sample capture-recapture model for an open population of animals. This model requires every animal to have both the same probability of capture in any sample and the same probability of survival for any period. For a detailed discussion of this model see Seber (1973; p. 196).

The Jolly-Seber model makes very strong assumptions so that there is a need for an omnibus goodness of fit test to it. Seber (1973; p. 223) suggests a traditional chi-square goodness of fit test based on comparing observed and expected values for numbers of animals captured with each possible capture history. He emphasizes that often many cells will need to be pooled because of their small expected values. Seber (1973; p. 224) also suggests an indirect test based on work of Leslie et al. (1953). This test compares newly marked individuals with estimates of this number obtained from animals captured at least twice. Jolly (1982) has also proposed an omnibus test. To avoid the large scale pooling necessary for Seber's chi-square test, he suggests comparing observed and expected values of the triangular array of animals caught in sample i which were last previously caught in sample h using a chi-square goodness of fit test.

Specific tests for individual assumptions have also been proposed. Carothers (1971) has suggested a specific test for heterogeneity of capture probabilities which extends earlier work of Leslie (1958). Robson (1969) and Pollock (1975) have suggested tests for survival and

capture probabilities being influenced by marking for a temporary period. (See also Brownie and Robson (1983)). Balser (unpublished 1981 Ph.D. thesis, Cornell University) has provided a test for temporary emigration in the Jolly-Seber model. His test is based on conditioning on sufficient statistics and turns out to be asymptotically equivalent to the omnibus test suggested by Jolly (1982).

Here we develop some conditional tests of the Jolly-Seber model. We then examine their power by conducting a detailed simulation study. Finally there is a general discussion section.

2. Some Definitions

Jolly (1965) and Seber (1965) in their formulation of the Jolly-Seber model assume binomial variation for the sample sizes and for the number surviving during any given period. Here, following Robson (1969) and Pollock (1975), we find it convenient to treat the sample sizes as fixed observable numbers. Also the number of marked animals at any period is treated as a fixed unknown parameter. However, our tests also apply to the original Jolly-Seber model. Allowance is also made for "losses on capture" by assuming that the number of animals released, after a sample has been taken, is a fixed observable number.

The following notation which is an obvious extension of that of Seber (1973) is used in this paper:

M_i, the number of marked animals in the population at the time the ith sample is taken ($i = 1, \ldots, k$; $M_1 \equiv 0$).

N_i, the total number of animals in the population at the time the ith sample is taken ($i = 1, \ldots, k$).

B_i, the total number of new animals entering the population between the ith and $(i + 1)$th sample. ($i = 1, \ldots, k-1$).

ϕ_i , the survival probability for all animals between the ith and (i + 1)th sample (i = 1, ..., k-1).

$X_{h_{i-1}1}$, the number of animals captured in the ith sample which have the previous capture history h_{i-1} (i = 1, ..., k).

$R_{h_{i-1}1}$, the number of the $X_{h_{i-1}1}$ animals which are released after the ith sample. Note that those animals which are unmarked are marked before being released. (i = 1, ..., k).

$m_i(u_i)$, the number of marked (unmarked) animals captured in the ith sample (i = 1, ..., k).

m_{hi} , the number of marked animals captured in the ith sample which were last previously captured in sample h (h = 1, ..., i-1; i = 2, ..., k).

n_i , $(m_i + u_i)$, the total number of animals captured in the ith sample (i = 1, ..., k).

R_i , the number of the n_i which are released after the ith sample (i = 1, ..., k-1).

Z_i , the number of animals captured before i, not captured at i, and captured again later (i = 2, ..., k-1).

r_i , the number of the R_i animals released at i which are recaptured again (i = 1, ..., k-1).

T_i , the number of marked animals in the population just before the ith sample which are recaptured (i = 1, ..., k-1).

3. Goodness of Fit Tests

3.1 The Likelihood Function

The joint probability distribution of the $X_{h_{i-1}1}$ which is the likelihood function can be expressed as

$$L = \Pr\left[\{X_{h_{k-1}1}\}, \{X_{h_{k-2}1}\}, \ldots, \{X_{h_1 1}\}\right]$$

$$= \Pr\left[m_2, \ldots, m_k, r_2, \ldots, r_{k-1}\right] \Pr\left[\{X_{h_{k-1}1}\}, \ldots, \{X_{h_1 1}\} \mid m_2 \ldots r_{k-1}\right]$$

The first component involves the minimal sufficient statistic which is of dimension (2k-3) and can be expressed in the following way.

$$\Pr\left[m_2, \ldots, m_k, r_2, \ldots, r_{k-1}\right]$$

$$= \prod_{i=2}^{k} \frac{\binom{M_i}{m_i}\binom{N_i - M_i}{n_i - m_i}}{\binom{N_i}{n_i}} \prod_{i=2}^{k-1} \frac{\binom{M_i - m_i}{Z_i}\binom{R_i}{r_i}}{\binom{M_i - m_i + R_i}{Z_i + r_i}}$$

The second component which is of dimension $(2^k - 3k+2)$ involves the joint probability distribution vectors of numbers in each capture history conditional on the minimal sufficient statistic. Clearly this component will not involve the unknown parameters (from the definition of sufficiency). We now present two different forms for this conditional distribution which form the basis of our goodness of fit tests.

3.2 One Conditional Goodness of Fit Test

The first form is given by

$$\Pr\left[\{X_{h_{k-1}1}\}, \ldots, \{X_{h_1}\} \mid m_2, \ldots, m_k, r_2, \ldots, r_{k-1}\right]$$

$$= \left\{ \prod_{i=2}^{k-1} \left[\prod_{h_{i-1}}^{*} \frac{\binom{R_{h_{i-1}1}}{X_{h_{i-1}11}, X_{h_{i-1}101}, \ldots, X_{h_{i-1}10\ldots01}}}{\binom{R_i}{r_i}} \right] \right\} \cdot$$

$$\left\{ \prod_{i=3}^{k} \frac{\binom{m_{1i}^{2i}}{m_{2i}}}{\binom{r_2+Z_2}{r_2}} \prod_{i=4}^{k} \frac{\binom{m_{1i}^{3i}}{m_{3i}}}{\binom{r_3+Z_3}{r_3}} \cdots \frac{\binom{m_{k-1}}{m_{k-2,k-1}}\binom{Z_{k-1}}{m_{k-2,k}}}{\binom{r_{k-2}+Z_{k-2}}{r_{k-2}}} \right\}$$

Note that $\prod_{h_{i-1}}^{*}$ denotes the product of all the capture histories. Also note that the notation in the second component needs elaboration. To illustrate we note

$$m_{1i}^{2i} = m_{1i} + m_{2i}$$

$$m_{1i}^{3i} = m_{1i} + m_{2i} + m_{3i}$$

It is interesting to note that the second term in the distribution which has dimension $\left(\frac{k(k-1)}{2} - 2k + 3\right)$ has been given by Balser in his Ph.D. thesis as a test for temporary emigration. It is also asymptotically equivalent to a test suggested by Jolly (1982). We shall refer to this as the Jolly-Balser test component.

3.3 Another Conditional Goodness of Fit Test

An alternative form is given by

$$\Pr\left[\{X_{h_{k-1}1}\}, \ldots, \{X_{h_1 1}\} \mid m_2, \ldots, m_k, r_2, \ldots, r_{k-1}\right]$$

$$= \left\{ \prod_{i=2}^{k-1} \left[\prod_{h_{i-1}}^{*} \frac{\binom{R_{h_{i-1}1}}{X_{h_{i-1}11} + X_{h_{i-1}101} + \ldots + X_{h_{i-1}10\ldots01}}}{\binom{R_i}{r_i}} \right] \right\} \cdot \left\{ \prod_{i=3}^{k-1} \left[\prod_{h_{i-1}}^{+} \frac{\binom{X_{h_{i-1}11} + X_{h_{i-1}01} + \ldots + X_{h_{i-1}0\ldots01}}{X_{h_{i-1}1}}}{\binom{T_{i-1}}{m_i}} \right] \right\}$$

$\prod_{h_{i-1}}^{*}$ denotes product over all capture histories

$\prod_{h_{i-1}}^{+}$ denotes product over all capture histories except {0, 0, ..., 0}

3.4 Illustration of Tests

In this section we present the form of the probability distributions for the special case when k = 5. We also show how the distributions give rise to contingency table chi-square tests. We believe that this section is very important to obtaining a clear understanding of the tests' structure.

3.4.1 The First Form

$$P\left[\{X_{h_4 1}\}, \{X_{h_3 1}\}, \{X_{h_2 1}\}, \{X_{h_1 1}\} \mid m_2, m_3, m_4, m_5, r_2, r_3, r_4\right]$$

$$= \left\{ \frac{\binom{R_{11}}{X_{111}, X_{1101}, X_{11001}} \binom{R_{01}}{X_{011}, X_{0101}, X_{01001}}}{\binom{R_2}{r_2}} \right.$$

$$\frac{\binom{R_{111}}{X_{1111}, X_{11101}} \binom{R_{101}}{X_{1011}, X_{10101}} \binom{R_{011}}{X_{0111}, X_{01101}} \binom{R_{001}}{X_{0011}, X_{00101}}}{\binom{R_3}{r_3}}$$

$$\frac{\binom{R_{1111}}{X_{11111}} \binom{R_{1011}}{X_{10111}} \binom{R_{0111}}{X_{01111}} \binom{R_{0011}}{X_{00111}} \binom{R_{1001}}{X_{10011}} \binom{R_{1101}}{X_{11011}} \binom{R_{0101}}{X_{01011}} \binom{R_{0001}}{X_{00011}}}{\binom{R_4}{r_4}}$$

$$\left. \left\{ \frac{\binom{m_{13}^{23}}{m_{23}} \binom{m_{14}^{24}}{m_{24}} \binom{m_{15}^{25}}{m_{25}}}{\binom{r_2 + Z_2}{r_2}} \quad \frac{\binom{m_{14}^{34}}{m_{34}} \binom{m_{15}^{35}}{m_{35}}}{\binom{r_3 + Z_3}{r_3}} \right\} \right.$$

These hypergeometric distributions give rise to the following independent contingency tables.

First Component (Degrees of Freedom 16)

i = 2 (Degrees of Freedom 3)

X_{111}	X_{1101}	X_{11001}	$R_{11} - X_{111} - X_{1101} - X_{11001}$
X_{011}	X_{0101}	X_{01001}	$R_{01} - X_{011} - X_{0101} - X_{01001}$

i = 3 (Degrees of Freedom 6)

X_{1111}	X_{11101}	$R_{111} - X_{1111} - X_{11101}$
X_{1011}	X_{10101}	$R_{101} - X_{1011} - X_{10101}$
X_{0111}	X_{01101}	$R_{011} - X_{0111} - X_{01101}$
X_{0011}	X_{00101}	$R_{001} - X_{0011} - X_{00101}$

i - 4 (Degrees of Freedom 7)

X_{11111}	$R_{1111} - X_{11111}$
X_{10111}	$R_{1011} - X_{10111}$
X_{01111}	$R_{0111} - X_{01111}$
X_{00111}	$R_{0011} - X_{00111}$
X_{10011}	$R_{1001} - X_{10011}$
X_{11011}	$R_{1101} - X_{11011}$
X_{01011}	$R_{0101} - X_{01011}$
X_{00011}	$R_{0001} - X_{00011}$

Second Component - Jolly-Balser (Degrees of Freedom 3)

Degrees of Freedom 2

m_{13}	m_{14}	m_{15}
m_{23}	m_{24}	m_{25}

Degrees of Freedom 1

$(m_{14} + m_{24})$	$(m_{15} + m_{25})$
m_{34}	m_{35}

3.4.2 The Second Form

$$P\left[\{X_{h_4 1}\}, \{X_{h_3 1}\}, \{X_{h_2 1}\}, \{X_{h_1 1}\} \mid m_2, m_3, m_4, m_5, r_2, r_3, r_4\right]$$

$$= \left\{ \frac{\binom{R_{11}}{X_{111} + X_{1101} + X_{11001}} \binom{R_{01}}{X_{011} + X_{0101} + X_{01001}}}{\binom{R_2}{r_2}} \right.$$

$$\frac{\binom{R_{111}}{X_{1111} + X_{11101}} \binom{R_{101}}{X_{1011} + X_{10101}} \binom{R_{011}}{X_{0111} + X_{01101}} \binom{R_{001}}{X_{0011} + X_{00101}}}{\binom{R_3}{r_3}}$$

$$\left. \frac{\binom{R_{1111}}{X_{11111}} \binom{R_{1011}}{X_{10111}} \binom{R_{0111}}{X_{01111}} \binom{R_{0011}}{X_{00111}} \binom{R_{1001}}{X_{10011}} \binom{R_{1101}}{X_{11011}} \binom{R_{0101}}{X_{01011}} \binom{R_{0001}}{X_{00011}}}{\binom{R_4}{r_4}} \right\}$$

$$\left\{ \frac{\binom{X_{101} + X_{1001} + X_{10001}}{X_{101}} \binom{X_{111} + X_{1101} + X_{11001}}{X_{111}} \binom{X_{011} + X_{0101} + X_{01001}}{X_{011}}}{\binom{T_2}{m_3}} \right.$$

$$\left. \frac{\binom{X_{1111}+X_{11101}}{X_{1111}} \binom{X_{1011}+X_{10101}}{X_{1011}} \binom{X_{0111}+X_{01101}}{X_{0111}} \binom{X_{0011}+X_{00101}}{X_{0011}} \binom{X_{1001}+X_{10001}}{X_{1001}} \binom{X_{1101}+X_{11001}}{X_{1101}} \binom{X_{0101}+X_{01001}}{X_{0101}}}{\binom{T_3}{m_4}} \right\}$$

This formulation of the conditional probability distribution gives rise to the following independent contingency tables.

First Component (Degrees of Freedom 11)

i = 2 (Degrees of Freedom 1)

$X_{111} + X_{1101} + X_{11001}$	$R_{11} - X_{111} - X_{1101} - X_{11001}$
$X_{011} + X_{0101} + X_{01001}$	$R_{01} - X_{011} - X_{0101} - X_{01001}$

i = 3 (Degrees of Freedom 3)

$X_{1111} + X_{11101}$	$R_{111} - X_{1111} - X_{11101}$
$X_{1011} + X_{10101}$	$R_{101} - X_{1011} - X_{10101}$
$X_{0011} + X_{00101}$	$R_{001} - X_{0011} - X_{00101}$
$X_{0011} + X_{00101}$	$R_{001} - X_{0011} - X_{00101}$

i = 4 (Degrees of Freedom 7)

X_{11111}	$R_{1111} - X_{11111}$
X_{10111}	$R_{1011} - X_{10111}$
X_{01111}	$R_{0111} - X_{01111}$
X_{00111}	$R_{0011} - X_{00111}$
X_{10011}	$R_{1001} - X_{10011}$
X_{11011}	$R_{1101} - X_{11011}$
X_{01011}	$R_{0101} - X_{01011}$
X_{00011}	$R_{0001} - X_{00011}$

Second Component (Degrees of Freedom 8)

i = 3 (Degrees of Freedom 2)

X_{101}	$X_{1001} + X_{10001}$
X_{111}	$X_{1101} + X_{11001}$
X_{011}	$X_{0101} + X_{01001}$

i = 4 (Degrees of Freedom 6)

X_{1111}	X_{11101}
X_{1011}	X_{10101}
X_{0111}	X_{01101}
X_{0011}	X_{00101}
X_{1001}	X_{10001}
X_{1101}	X_{11001}
X_{0101}	X_{01001}

3.5 Practical Application of the Tests (Pooling Suggestions)

The full contingency tables for each of the 2 tests contain too many cells for practical application with most actual data sets, and pooling of cells is thus generally necessary. Cell pooling decisions can always be based on visual inspection of contingency tables, but we were interested in standardized pooling rules that could be implemented in computer programs.

The first test has 2 components, the first of which includes k-2 tables. The rows of these tables correspond to the different previous capture histories of animals captured in period i (i = 2, ..., k-1). We suggest reducing the number of rows to 2, with one row corresponding to animals captured for the first time at period i and the other row corresponding to all animals captured both at i and some previous period (< i). The columns of these k-2 tables correspond to the periods in which the animals captured in i are next captured (i + 1, i + 2, ...k), with a final column for animals never captured again after i. We recommend that the tables for i = 2, 3, ...k-2 be collapsed to yield 3 columns, one for animals next captured at i + 1, one for animals next captured in some period > i + 1 (this column is thus obtained by pooling) and one for animals never recaptured after i. For the last table, corresponding to period k - 1, only 2 columns are possible; animals captured at k and animals not recaptured. The second (Jolly-Balser) component of the first test involves the m_{ij} statistics, and we chose not to apply an initial pooling rule to the k-3 resulting contingency tables.

The second test also has 2 components, the first of which includes k-2 contingency tables. The rows of these tables correspond to the different previous capture histories of animals captured in period i (i = 2, ...k-1). We recommend pooling all animals captured both at i and some previous period (< i) to form one row, and using animals captured for the first time at i to form the other row. The 2 columns of these tables correspond to animals that are and are not captured again at some period > i, and these, of course, require no pooling. Note that with these pooling rules, this first component of the second test reduces to the test for short term mortality

due to marking given by Robson (1969) and Brownie and Robson (1983). The second component of the second test includes k-3 contingency tables. The rows of the tables correspond to the different previous capture histories of animals captured at i (i = 2, ...k-2) and at some later period (> i). We recommend pooling all of these animals that were captured at some period < i to form one row, and using all animals captured for the first time at i to form a second row. The 2 columns of these tables correspond to animals next recaptured at i + 1 and to animals next recaptured at some later period (> i + 1).

In addition to this initial pooling, which we implement for all data sets, it is sometimes necessary to pool additional cells when expected values are small. Our approach has been to pool cells with the smallest expectations until all expected values are ≥ 2.

4. Simulation Study of Power of the Tests

4.1 Description

We used computer simulation to investigate the size and power of the 2 suggested tests against alternatives involving heterogeneity of capture and survival probabilities. We recognized 2 subpopulations of animals in the simulation program. The first subpopulation (denoted by primed symbols) was started with $N_1' = \pi N_1$ individuals and the second subpopulation (denoted by double-primed symbols) was started with $N_1'' = (1 - \pi)N_1$ individuals, where π denotes the proportion of the total initial population in the first subgroup. Capture and survival of each individual for each sampling period were treated as independent Bernoulli

trials using pseudorandom numbers. Although our simulation program was quite general, we made the following restrictive assumptions to facilitate interpretation of results:

$$\phi_i' = \phi' \text{ for all } i,$$

$$\phi_i'' = \phi'' \text{ for all } i,$$

$$p_i' = p' \text{ for all } i, \text{ and}$$

$$p_i'' = p'' \text{ for all } i.$$

We also assumed no losses on capture. Constant numbers of new individuals were added to each subpopulation every time period after the first ($B_i' = B' = (1-\phi')N_1'$, $B_i'' = B'' = (1-\phi'')N_1''$ for $i = 1, \ldots, k-1$). Since the number of births balanced expected deaths, expected sizes of the subpopulations were constant, $E(N_i') = N_i'$ and $E(N_i'') = N_1''$, for $i = 1, \ldots, k-1$.

The above procedure was used to generate a table of capture histories for the entire population each iteration. Summary statistics for these capture histories were then used to compute test statistics using the 2 described goodness-of-fit tests with the pooling modifications given in Section 3.5. For each component of each test, the resulting statistic was compared with the critical values corresponding to the 0.01, 0.05 and 0.10 probability levels. Power was then estimated for each probability level as b/c, where b is the total number of rejections recorded among the c iterations. The computer program itself was written in standard FORTRAN for an HP3000 Series 3 computer and is available from the second author, James E. Hines.

4.2 Heterogeneity of Capture Probabilities

To illustrate the dependence of the power of the test on population size (N), sample size (k) and survival rate (ϕ) when capture probabilities are heterogeneous we present Figure 1. This figure is for the overall test using the first formulation but trends are similar for the second formulation. Notice that the power increases with sample size, population size and survival rate as you would expect. Also notice that for the parameters in many practical applications the test will have low power.

(Figure 1 to appear here)

We also did a comparison of the two test formulations and the components of each test (Table 1). Notice that test 1 performs somewhat better than test 2 under heterogeneity of capture probabilities. Also notice that here the components of the tests behave similarly.

(Table 1 to appear here)

4.3 Heterogeneity of Survival Probabilities

To illustrate the dependence of the power of the test on population size (N), sample size (k) and capture probability (p) when survival probabilities are heterogeneous we present Figure 2. This figure is for the overall test using the first formulation but again trends are similar for the second formulation. Notice that the power increases with sample size, population size and capture probability as you would expect. Again it is true that for the parameters of many practical applications the test will have low power.

(Figure 2 to appear here)

We also did a comparison of the two test formulations and the components of each test (Table 2). Here test 2 performs a little better than test 1 which is the opposite of that in § 4.2. For test 1 the first component provides all the power. The second component which is the Jolly-Balser has no power against heterogeneity of survival probabilities. For the test 2 both components have some power but component 2 is much more powerful than component 1.

(Table 2 to appear here)

4.4 Heterogeneity of Capture and Survival Probabilities

We also did a small comparison of the two tests when both capture and survival probabilities were heterogeneous (Table 3). We found that there was little to choose between the two tests overall and that they both performed much better when there was a positive relationship between the subgroup capture and survival probabilities. The results for the components of the tests were extremely interesting. For test 1 the first component performed better for a direct relationship (like the overall test) while the second component (Jolly-Balser) performed better for an inverse relationship. For test 2 the results were reversed. The first component performed best for the inverse relationship while the second performed best for the direct relationship.

(Table 3 to appear here)

5. Discussion

In this paper we have presented some goodness of fit tests to the Jolly-Seber model. We have examined their power properties for heterogeneity of capture, survival and a combination. Our results although condensed show the need for examination of the test components as well as the overall tests. Our results also show that often in practice the tests will have low power.

We did not examine the important alternative of temporary emigration. We felt this inappropriate as Balser examined this in his thesis which is still unpublished.

We wish to emphasize that goodness of fit tests cannot detect all types of failures of assumptions. Permanent trap response of capture probabilities cannot be detected. Also a permanent lowering of survival due to marking, an important problem in fisheries, cannot be detected. However it is possible to detect short term influences on survival or capture probabilities (Robson 1969, Pollock 1975, Brownie and Robson 1983). Finally we also mention that mark loss will be very difficult to detect by goodness of fit tests. If the mark loss is a function of time since marking the test may have some power to detect it.

REFERENCES

Brownie, C. and Robson, D. S. (1983). Estimation of time-specific survival rates from tag-resighting samples: A generalization of the Jolly-Seber model. Biometrics 39, 000-000.

Carothers, A. D. (1971). An examination and extension of Leslie's test of equal catchability. Biometrics 27, 615-630.

Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika 52, 225-247.

Jolly, G. M. (1982). Mark-recapture models with parameters constant in time. Biometrics 38, 301-321.

Leslie, P. H. (1958). Statistical appendix. J. Animal Ecol. 27, 84-86.

Leslie, P. H., Chitty, D. and Chitty, H. (1953). The estimation of population parameters from data obtained by means of the capture-recapture method III. An example of the practical applications of the method. Biometrika 40, 137-169.

Pollock, K. H. (1975). A K-sample tag-recapture model allowing for unequal survival and catchability. Biometrika 62, 577-583.

Pollock, K. H. (1981). Capture-recapture models allowing for age-dependent survival and capture rates. Biometrics 37, 521-529.

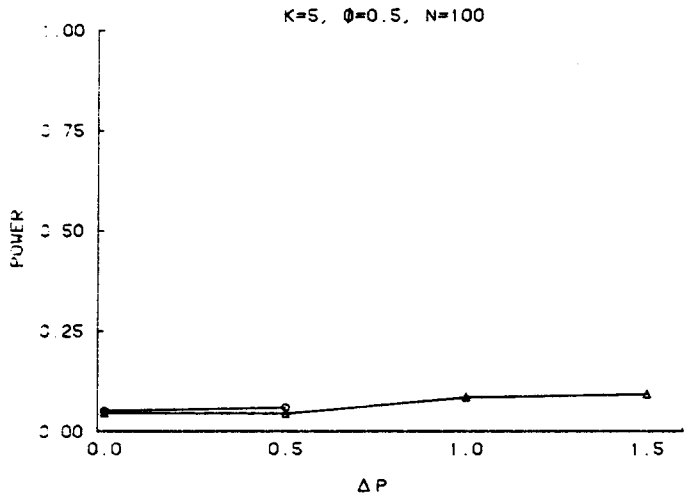
Robson, D. S. (1969). Mark-recapture methods of population estimation. In New Developments in Survey Sampling, N. L. Johnson and H. Smith, Jr. (eds.), 120-140. New York, New York: Wiley-Interscience.

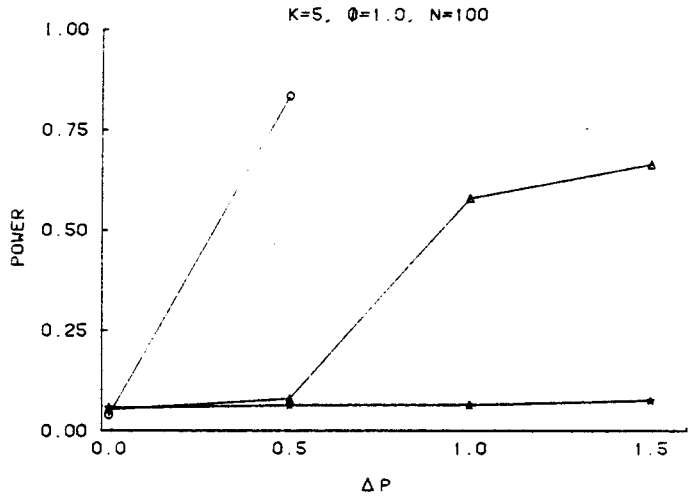
Seber, G. A. F. (1965). A note on the multiple-recapture census. Biometrika 52, 249-259.

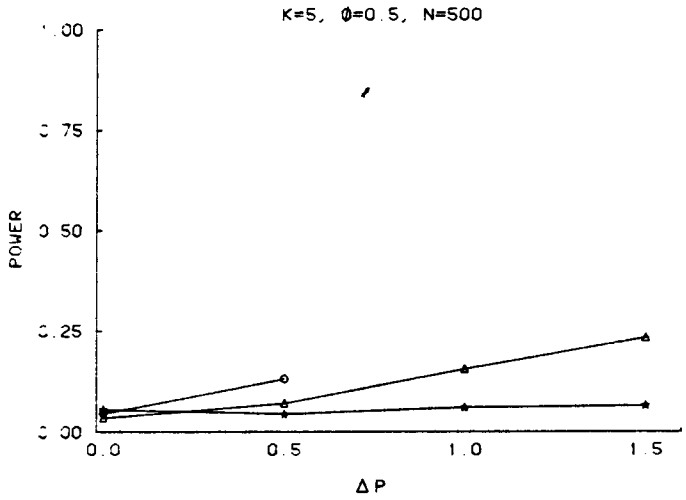
Seber, G. A. F. (1973). The Estimation of Animal Abundance and Related Parameters. London: Griffin.

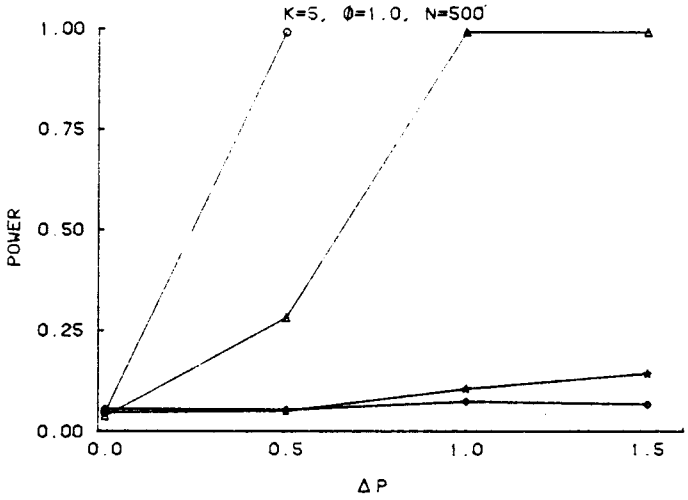
FIGURE LEGEND

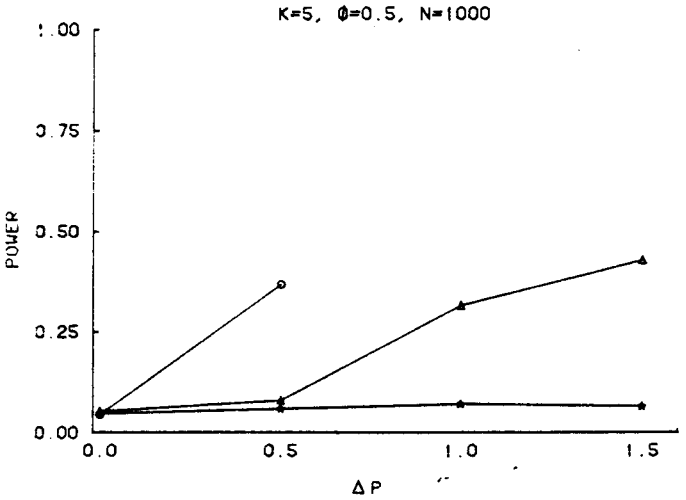
Figure 1 - Empirical power estimates for the first goodness of fit test for differing values of survival (ϕ) and population size (N) as a function of increasing heterogeneity of capture probabilities. We define $\Delta p = (p' - p'')/E(p)$ and use $\Diamond$ when $E(p) = 0.1$, * when $E(p) = 0.2$, Δ when $E(p) = 0.5$ and 0 when $E(p) = 0.8$. Part (a) refers to sampling periods $k = 5$ and part (b) to sampling periods $k = 10$. The size of test is $\alpha = 0.05$.

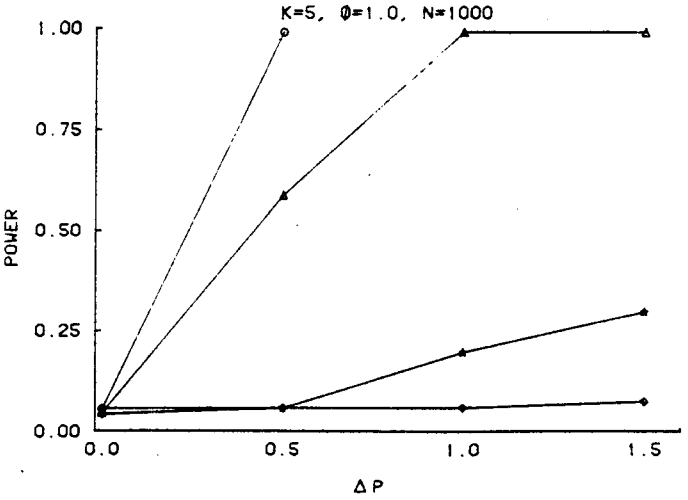

K=5, Ø=0.5, N=100
POWER
ΔP
0.0
0.5
1.0
1.5

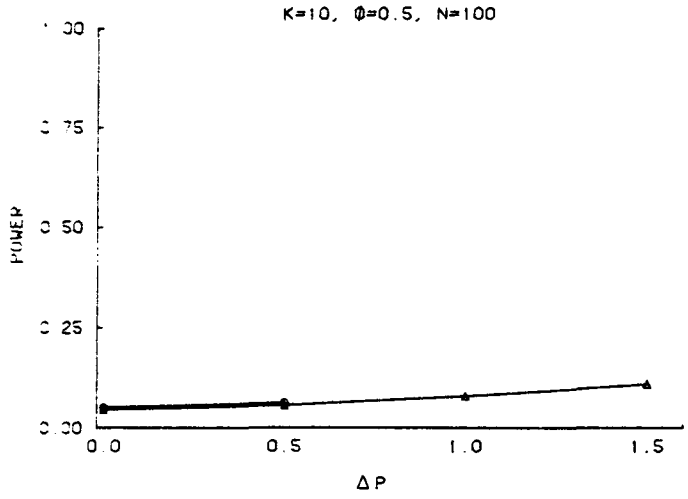

K=5, Ø=1.0, N=100
POWER
1.00
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
ΔP

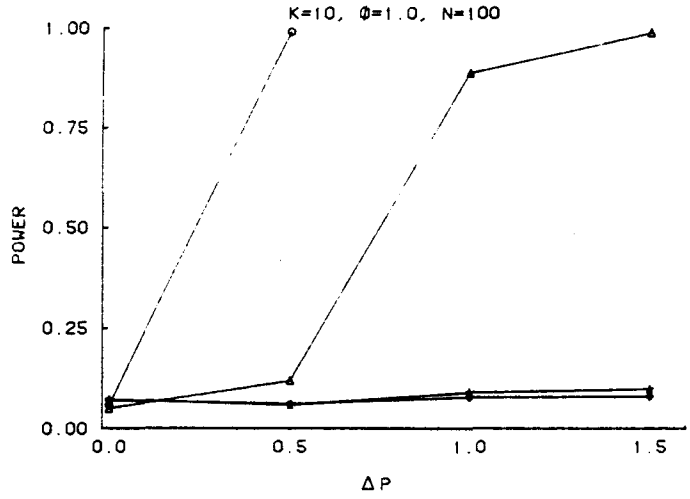

K=5, Ø=0.5, N=500
POWER
ΔP
0.0
0.5
1.0
1.5

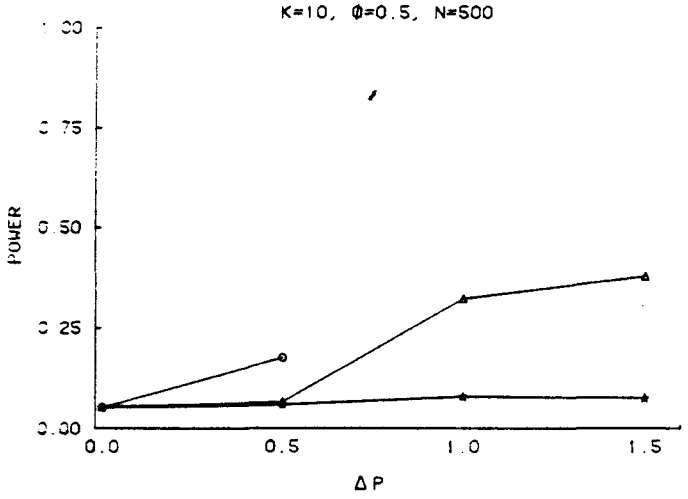

K=5, Ø=1.0, N=500
POWER
1.00
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
ΔP

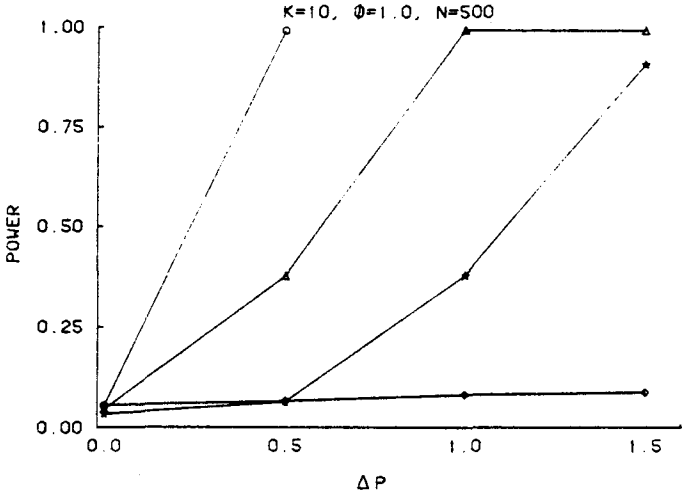

K=5, Ø=0.5, N=1000
POWER
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
ΔP

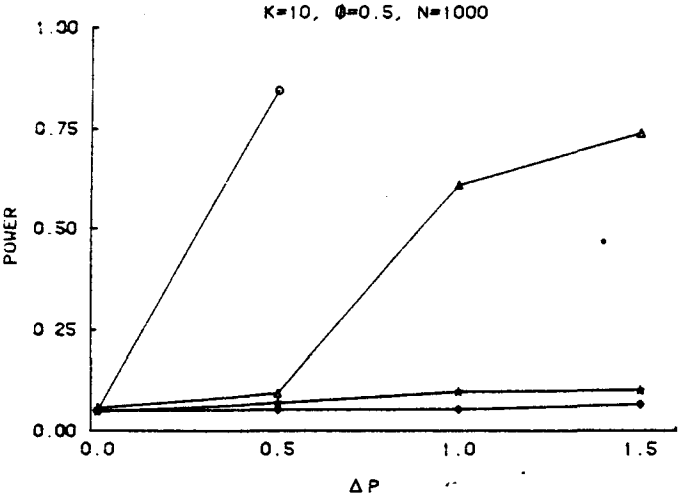

K=5, Ø=1.0, N=1000
POWER
1.00
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
ΔP


K=10, Φ=0.5, N=100
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
ΔP


K=10, Φ=1.0, N=100
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
ΔP


K=10, Φ=0.5, N=500
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
ΔP


K=10, Φ=1.0, N=500
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
ΔP


K=10, Φ=0.5, N=1000
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
ΔP

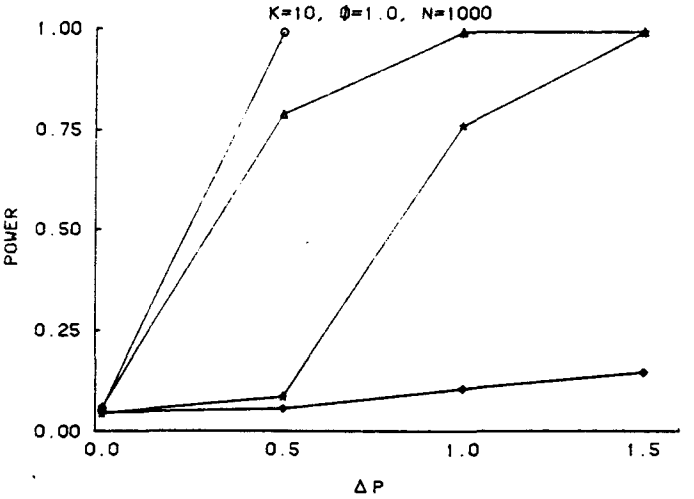

K=10, Φ=1.0, N=1000
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
ΔP

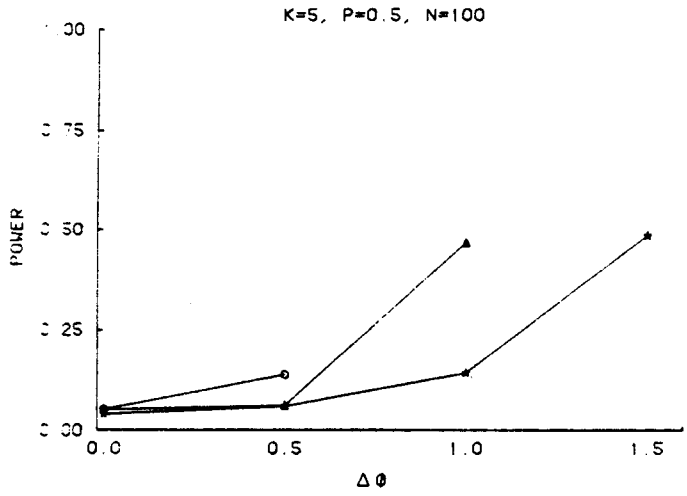
Table 1. Empirical power estimates for both goodness of fit tests for $N = 500$, $K = 10$ and p_1, p_2, ϕ_1, ϕ_2 as specified. Heterogeneity of capture probabilities.[a] Size of test $\alpha = 0.05$.

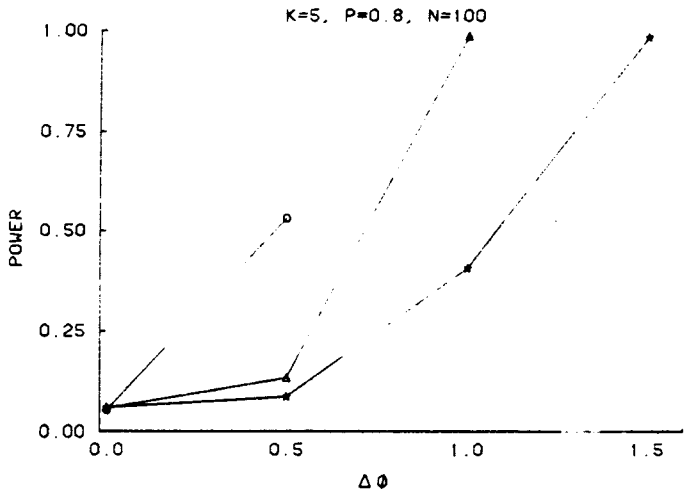
p_1	p_2	ϕ_1	ϕ_2	Test 1			Test 2		
				Comp 1	Comp 2	Total	Comp 1	Comp 2	Total
0.15	0.25	0.50	0.50	0.0690	0.0570	0.0610	0.0650	0.0700	0.0680
0.15	0.25	1.00	1.00	0.0820	0.0500	0.0650	0.0550	0.0790	0.0810
0.375	0.625	0.50	0.50	0.0710	0.0650	0.0670	0.0630	0.0530	0.0650
0.375	0.625	1.00	1.00	0.3220	0.2510	0.3770	0.1370	0.0942	0.1420
0.60	1.00	0.50	0.50	0.1570	0.0490	0.1770	0.0460	0.0930	0.0710
0.60	1.00	1.00	1.00	0.9980	1.0000	1.0000	0.9910	-	0.9910

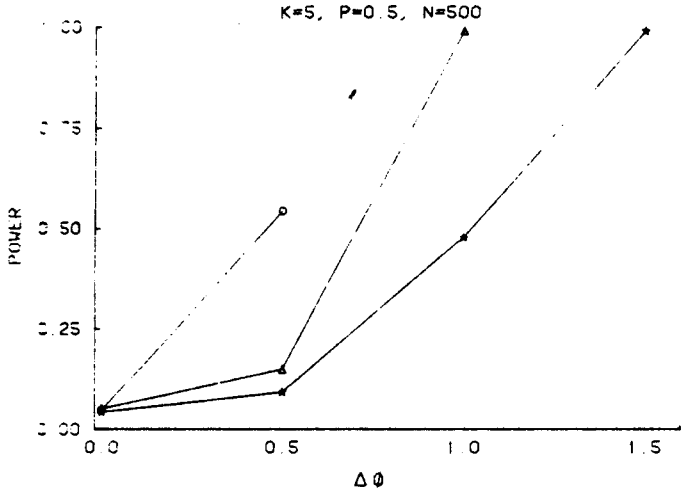
[a] Power was estimated for each component of each test separately, as well as for the entire tests.

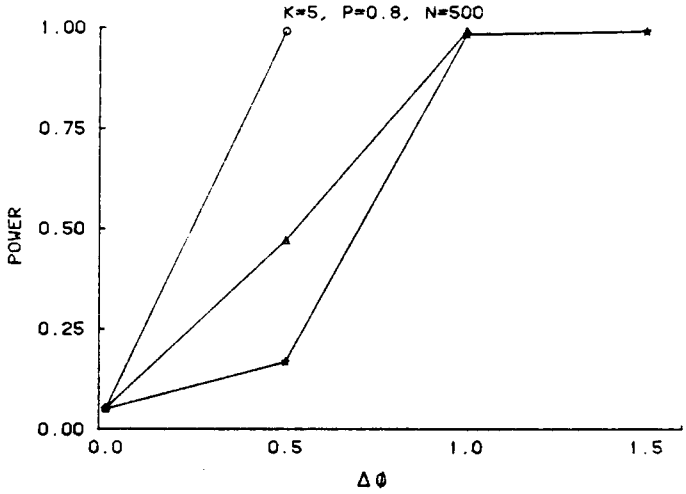
FIGURE LEGEND

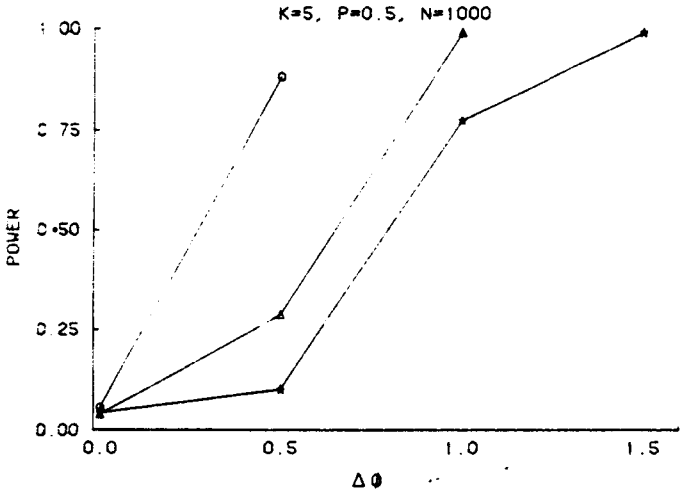
Figure 2 - Empirical power estimates for the first goodness of fit test for differing values of capture probability (p) and population size (N) as a function of increasing heterogeneity of survival probabilities. We define $\Delta\phi = (\phi' - \phi'')/E(\phi)$ and use * when $E(\phi) = 0.2$, Δ when $E(\phi) = 0.5$, and 0 when $E(\phi) = 0.8$. Part (a) refers to sampling periods $k = 5$ and part (b) to sampling periods $k = 10$. The size of test is $\alpha = 0.05$.

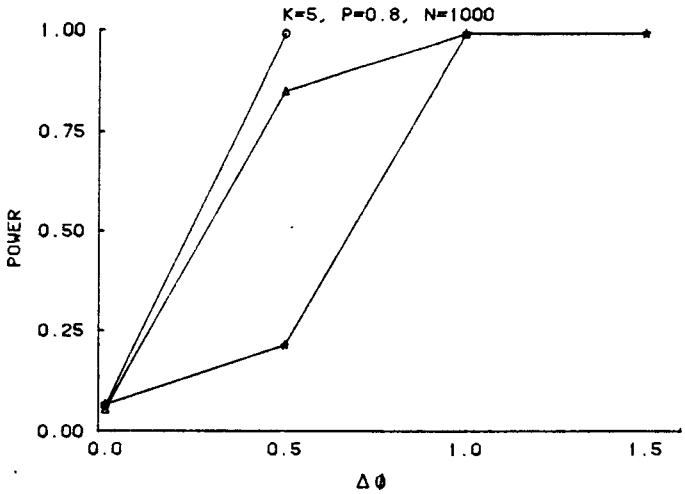

K=5, P=0.5, N=100
POWER
0.00
0.25
0.0
0.5
1.0
1.5
ΔØ

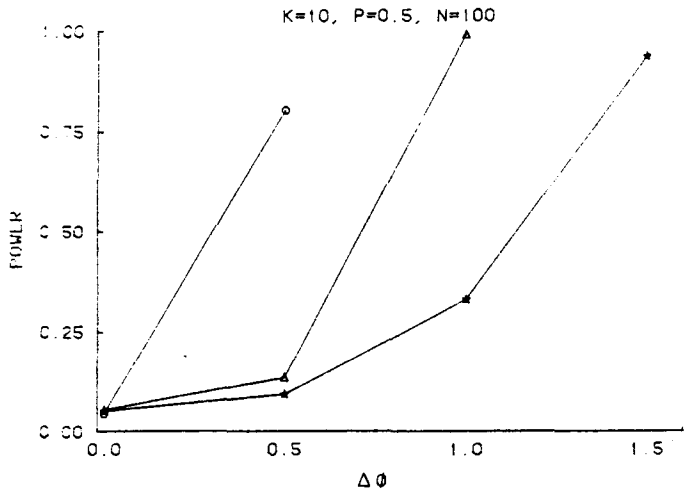

K=5, P=0.8, N=100
POWER
1.00
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
ΔØ

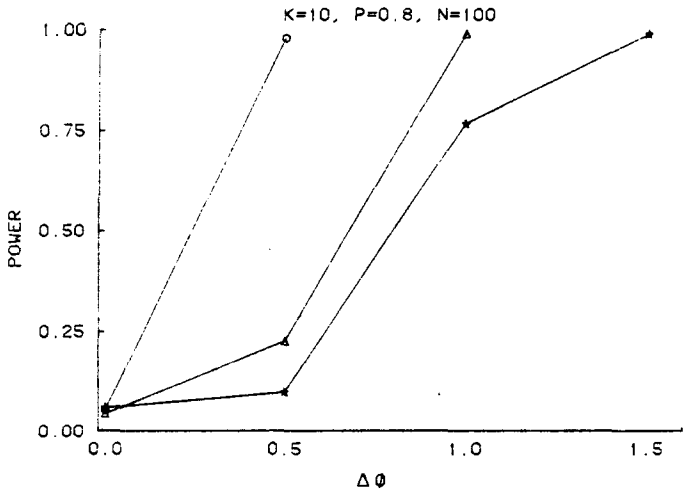

K=5, P=0.5, N=500
POWER
0.25
0.0
0.5
1.0
1.5
ΔØ

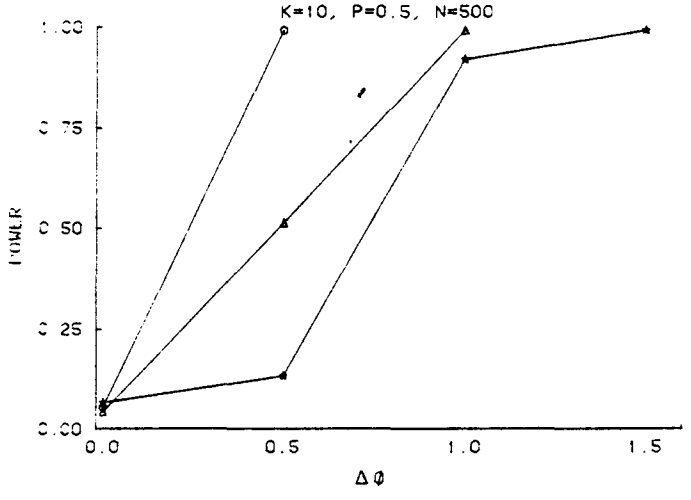

K=5, P=0.8, N=500
POWER
1.00
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
ΔØ

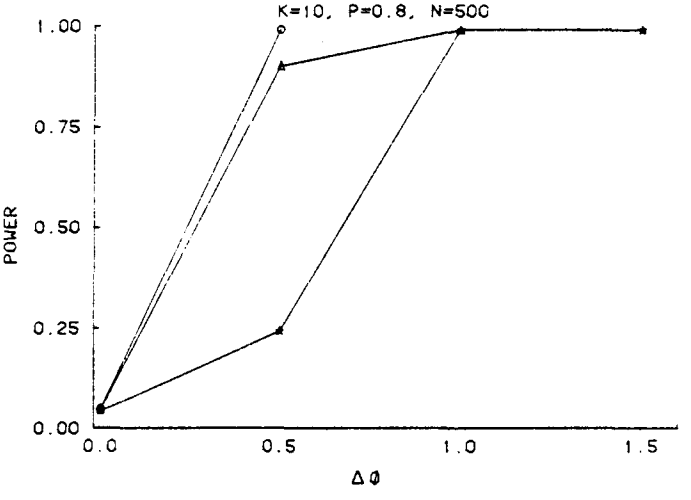

K=5, P=0.5, N=1000
POWER
0.50
0.25
0.00
0.0
0.5
1.0
1.5
ΔØ

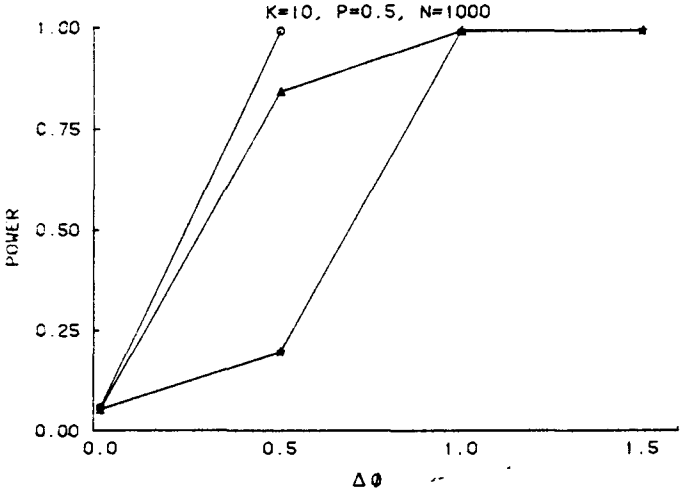

K=5, P=0.8, N=1000
POWER
1.00
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
ΔØ


K=10, P=0.5, N=100
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
Δ∅


K=10, P=0.8, N=100
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
Δ∅


K=10, P=0.5, N=500
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
Δ∅


K=10, P=0.8, N=500
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
Δ∅


K=10, P=0.5, N=1000
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
Δ∅

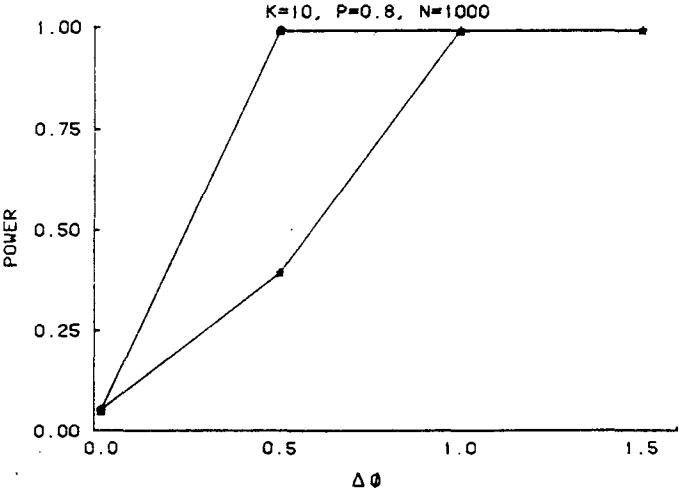

K=10, P=0.8, N=1000
POWER
0.00
0.25
0.50
0.75
1.00
0.0
0.5
1.0
1.5
Δ∅

Table 2. Empirical power estimates for both goodness of fit tests for $N = 500$, $K = 10$ and p_1, p_2, ϕ_1, ϕ_2 as specified. Heterogeneity of Survival probabilities. [a] Size of test $\alpha = 0.05$.

				Test 1			Test 2		
p_1	p_2	ϕ_1	ϕ_2	Comp 1	Comp 2	Total	Comp 1	Comp 2	Total
0.2	0.2	0.45	0.75	0.1610	0.0600	0.1370	0.0640	0.1600	0.1530
0.5	0.5	0.45	0.75	0.6100	0.0860	0.5120	0.1170	0.6970	0.6400
0.8	0.8	0.45	0.75	0.9450	0.0740	0.9010	0.0810	0.9710	0.9450
0.2	0.2	0.60	1.00	0.6700	0.0610	0.4420	0.1040	0.7560	0.6750
0.5	0.5	0.60	1.00	1.0000	0.0760	1.0000	0.5170	1.0000	1.0000
0.8	0.8	0.60	1.00	1.0000	0.0730	1.0000	0.3350	1.0000	1.0000

[a] Power was estimated for each component of each test separately, as well as for the entire tests.

Table 3. Empirical power estimates for both goodness-of-fit tests for $N = 500$, $K = 10$, and p_1, p_2, ϕ_1, ϕ_2 as specified. Heterogeneity of capture and survival probabilities.[a] Size of test $\alpha = 0.05$.

				Test 1			Test 2		
p_1	p_2	ϕ_1	ϕ_2	Comp 1	Comp 2	Total	Comp 1	Comp 2	Total
0.15	0.25	0.30	0.50	0.1287	0.0566	0.0997	0.0645	0.1241	0.1045
0.15	0.25	0.50	0.30	0.0538	0.0675	0.0711	0.0646	0.0520	0.0579
0.15	0.25	0.45	0.75	0.3130	0.0540	0.2380	0.0460	0.3590	0.2820
0.15	0.25	0.75	0.45	0.0620	0.0850	0.1010	0.0800	0.0640	0.0810
0.15	0.25	0.60	1.00	0.9940	0.0510	0.9460	0.0640	0.9990	0.9960
0.15	0.25	1.00	0.60	0.1310	0.1050	0.1520	0.1880	0.1190	0.2100
0.375	0.625	0.30	0.50	0.4530	0.0370	0.3880	0.0510	0.5370	0.4210
0.375	0.625	0.50	0.30	0.0590	0.1523	0.1210	0.1130	0.0570	0.0960
0.375	0.625	0.45	0.75	0.9990	0.0500	0.9910	0.0500	0.9990	0.9990
0.375	0.625	0.75	0.45	0.1100	0.4110	0.3650	0.3440	0.0740	0.2730
0.375	0.625	0.60	1.00	1.0000	0.0450	1.0000	0.0600	1.0000	1.0000
0.375	0.625	1.00	0.60	0.9920	0.8100	0.9990	0.9870	0.9340	1.0000
0.60	1.00	0.30	0.50	0.9640	-	0.9640	0.1324	0.9550	0.9420
0.60	1.00	0.50	0.30	0.0700	0.5337	0.1410	0.2730	0.0550	0.2010
0.60	1.00	0.45	0.75	1.0000	-	1.0000	0.9200	1.0000	1.0000
0.60	1.00	0.75	0.45	0.3870	1.0000	0.9980	0.9330	0.1370	0.8750
0.60	1.00	0.60	1.00	1.0000	0.9535	1.0000	1.0000	1.0000	1.0000
0.60	1.00	1.00	0.60	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

[a]Power was estimated for each component of each test separately, as well as for the entire tests.